(2)

ETL-0543

Vision-Based Navigation for Autonomous Ground Vehicles

Summary Report

Larry S. Davis

Center for Automation Research
University of Maryland
College Park, Maryland 20742-3411

August 1989

Approved for public release; distribution is unlimited.

Prepared for:

Defense Advanced Research Projects Agency
1400 Wilson Boulevard
Arlington, Virginia 22209-2308

U.S. Army Corps of Engineers
Engineer Topographic Laboratories
Fort Belvoir, Virginia 22060-5546

UNCLASSIFIED
SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1a. REPORT SECURITY CLASSIFICATION
UNCLASSIFIED

1b. RESTRICTIVE MARKINGS

2a. SECURITY CLASSIFICATION AUTHORITY

2b. DECLASSIFICATION / DOWNGRADING SCHEDULE

3. DISTRIBUTION / AVAILABILITY OF REPORT
Approved for public release;
distribution is unlimited.

4. PERFORMING ORGANIZATION REPORT NUMBER(S)

5. MONITORING ORGANIZATION REPORT NUMBER(S)
ETL-0543

6a. NAME OF PERFORMING ORGANIZATION
University of Maryland

6b. OFFICE SYMBOL (If applicable)

7a. NAME OF MONITORING ORGANIZATION
U.S. Army Engineer Topographic Laboratories

6c. ADDRESS (City, State, and ZIP Code)
Center for Automation Research
College Park, MD 20742-3411

7b. ADDRESS (City, State, and ZIP Code)
Fort Belvoir, VA 22060-5546

8a. NAME OF FUNDING / SPONSORING ORGANIZATION
Defense Advanced Research Projects Agency

8b. OFFICE SYMBOL (If applicable)
ISTO

9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER
DACA76-84-C-0004

8c. ADDRESS (City, State, and ZIP Code)
1400 Wilson Boulevard
Arlington, VA 22209-2308

10. SOURCE OF FUNDING NUMBERS

PROGRAM ELEMENT NO.	PROJECT NO.	TASK NO.	WORK UNIT ACCESSION NO.
62301E			

11. TITLE (Include Security Classification)
Vision-Based Navigation for Autonomous Ground Vehicles Summary Report

12. PERSONAL AUTHOR(S)
Larry S. Davis

13a. TYPE OF REPORT
Summary

13b. TIME COVERED
FROM 7/84 TO 12/87

14. DATE OF REPORT (Year, Month, Day)
August 1989

15. PAGE COUNT
133

16. SUPPLEMENTARY NOTATION
Previous reports in series: ETL-0509 (First Annual); ETL-0432 (1986 Annual); ETL-0479 (two-year summary); ETL-0519 (Third Annual)

17. COSATI CODES

FIELD	GROUP	SUB-GROUP

18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
Autonomous Navigation,
Road Following.
Computer Vision

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

This is a summary report for contract DACA76-84-C-0004, "Vision-Based Navigation for Autonomous Ground Vehicles." Our research has resulted in seventeen technical reports (list appended to this report, with abstracts), many of which have been subsequently published in journals, conferences and workshops. Additionally, our project involved close collaboration with the Martin Marietta Corporation, Denver, Colorado, in the development and testing of vision algorithms for navigation of roads and road networks. Several experiments were run on the Martin Autonomous Land Vehicle using programs developed at the University of Maryland, and some critical components of Martin Marietts's visula navigation system were based on fundamental research conducted at the University of Maryland under support of this contract--specifically, the overall framework of a focus-of-attention vision system, in which detailed analyses are performed on selected windows of images of roads, and the shape-from-contour algorithms (e.g., the zero-bank algorithm) that allowed the vehicle software to recover an accurate three-dimensional road model from monocular imagery, thus saving the autonomous land vehicle (ALV) from having to perform costly, and less reliable, analyses based on either stereo or motion.

20. DISTRIBUTION / AVAILABILITY OF ABSTRACT
☐ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT. ☐ DTIC USERS

21. ABSTRACT SECURITY CLASSIFICATION
UNCLASSIFIED

22a. NAME OF RESPONSIBLE INDIVIDUAL
Linda H. Graff

22b. TELEPHONE (Include Area Code)
(202) 355-2700

22c. OFFICE SYMBOL
CEETL-RI

DD Form 1473, JUN 86 Previous editions are obsolete.

SECURITY CLASSIFICATION OF THIS PAGE
UNCLASSIFIED

Preface

This report describes work performed under contract DACA76-84-C-0004 by the University of Maryland, College Park, Maryland, for the U.S. Army Engineer Topographic Laboratories (ETL), Fort Belvoir, Virginia, and the Defense Advanced Research Projects Agency (DARPA), Arlington, Virginia. The Contracting Officer's Representative at ETL is Ms. Linda Graff. The DARPA point of contact is Dr. William Isler.

Table of Contents

1. Introduction

This is a summary report for Contract DACA76-84-C-0004, Vision-based Navigation for Autonomous Ground Vehicles. Our research has resulted in seventeen technical reports (list appended to this report, with abstracts), many of which have been subsequently published in journals, conferences and workshops. Additionally, our project involved close collaboration with the Martin Marietta Corporation, Denver, Colorado, in the development and testing of vision algorithms for navigation of roads and road networks. Several experiments were run on the Martin Autonomous Land Vehicle using programs developed at the University of Maryland, and some critical components of Martin Marietta's visual navigation system were based on fundamental research conducted at the University of Maryland under support of this contract—specifically, the overall framework of a focus-of-attention vision system, in which detailed analyses are performed on selected windows of images of roads, and the shape-from-contour algorithms (e.g., the zero-bank algorithm) that allowed the vehicle software to recover an accurate three dimensional road model from monocular imagery, thus saving the ALV from having to perform costly, and less reliable, analyses based on either stereo or motion.

Our research has been extensively documented in 17 technical reports [1–17] and three annual reports [18–20]. In this summary report we will provide brief descriptions of the key technical contributions of the project, giving references to the technical reports in which more detailed explanations and examples can be found. We have performed research in three basic areas:

1) algorithms for visual navigation,
2) support of Martin Marietta in achieving program demonstration milestones, and
3) parallel implementations of algorithms for visual navigation.

We describe these three areas in the following subsections.

2. Algorithms for Visual Navigation

During the first year of the contract we designed and constructed an initial implementation of a visual navigation system for road network navigation. That system was based on a "focus of attention" vision principle. It developed a three-dimensional model of the road in front of the vehicle and used this model to predict the image locations of important road features, such as road boundaries and markings. By identifying the locations of these features the system was able both to verify its current three dimensional road model and to extend that model out further towards the horizon.

The system had modules for image processing, road geometry reconstruction, camera control, planning and navigation. Their activities were coordinated by a vision executive that controlled the flow of information and control between modules of the system. Early in the second year of the contract the system was able to routinely drive a robot arm carrying a black and white television camera over a terrain board, with some modest topography, on which we had painted a simple network of roads. The details of this initial road following system were described in technical reports [1,2,6].

It became clear to us early in the second year of the contract that the implementation methods used to construct that system made it somewhat inflexible and difficult to change. We were interested in being able to experiment with new control strategies for road detection and following, and this was not easily accomplished using the relatively rigid structure of our system. We therefore designed a production system version of the navigation system. This system was implemented using a set of communicating production systems coordinated through a structured blackboard. It was described in detail in technical report [15], which was recently accepted for publication in the IEEE Transactions on Robotics and Automation. We used the system to experiment with new road boundary detection strategies. For example, our initial system would track a straight road by sequentially placing overlapping windows on the image that were predicted to contain the road boundary; specialized image processing algorithms would then be applied to those windows to find the road boundaries. But if previous image analysis has revealed that the road is straight, then it would seem possible to track the road by placing windows that would "skip over" large parts of the straight road boundaries. With the system described in [15] it was relatively easy to describe such a strategy to the system, and to then study its behavior on a variety of road imagery.

During the second year of the contract, the ALV was equipped with an ERIM laser range scanner. Its main purpose was to allow the ALV to detect obstacles and navigate around those obstacles. The vision group at Martin Marietta developed an obstacle detection algorithm that was based on transforming the laser range data, recovered initially in a cylindrical coordinate system centered at the sensor, to a Cartesian coordinate system with z-coordinate corresponding to elevation above the ground. Their algorithms then marked as an obstacle any point whose elevation was above threshold.

We realized that the success of this algorithm depended on an accurate measurement of the attitude of the range scanner with respect to the ground, and that even small errors in the estimation of scanner attitude would lead to unacceptably large errors in the estimated elevation of pixels. We developed a more robust, alternate strategy for road obstacle detection based on comparing derivatives of observed range data against the predicted derivatives for a horizontal road. While this also required estimating the attitude of the scanner with respect to the road, we were able to show that our algorithm was far less sensitive to errors in this measurement process than the Martin Marietta algorithm. A set of comparative experiments were conducted both on synthetic data and on range images acquired from Martin Marietta. The results of this research were described in technical reports [13,14].

To support our research program in range data analysis, we designed and constructed a structured light range scanner. The scanner is described in technical report [17]. The scanner was small enough that it could be carried by our robot over the terrain board. Since it used the same TV camera that was used for road detection, we were able to acquire range data that is registered with the black and white video data. The range scanner was used by Prof. Minoru Asada of Osaka University, who spent one year visiting our Laboratory and working on the contract. In technical report [16] Prof. Asada described a set of algorithms that could fuse range data taken as the sensor moved through the world, thus developing a more complete and accurate map of the vehicle's environment.

One of the most important modules in both our system and Martin's system was the module that reconstructed the three dimensional geometry of the road. It was important that this be done accurately, because at the speeds that the vehicle was moving it could traverse over 20 feet between taking successive frames. The most obvious methods for road reconstruction—stereo and motion—were ruled out because of their sensitivity to certain calibration errors. This sensitivity was analyzed in technical report [11], where we showed that for reasonable values for the accuracy of estimation of vehicle heading, vehicle speed and vehicle position, the three dimensional locations of road boundaries recovered by time-varying image analysis algorithms would be far too inaccurate to be used to navigate the vehicle. Instead, both Martin Marietta and the University of Maryland studied the possibility of using monocular inverse perspective methods for road reconstruction.

The simplest such technique is the flat road model. If one assumes that the road is flat, and that one can measure the attitude and height of the camera with respect to the road, then one can determine the three dimensional location of any image point very simply. This method was used by Martin in early demonstrations. However, at Maryland we showed that this was not a very good approach for two reasons. First, it was very sensitive to errors in attitude estimation (similar to the elevation-based obstacle detection algorithm), and second, even small deviations of the road from flatness led to gross errors in the three-dimensional reconstruction. This latter problem was introduced by the relatively low grazing angle of the camera axis with respect to the road.

We developed a set of more sophisticated road inverse perspective algorithms, the most successful of which is the so-called zero bank algorithm. This algorithm is described in technical report [4]. It assumes that the elevation of the road may change, but that the road does not bank. similar to the geometry of a railroad track. Comparative experiments with this algorithm and other monocular inverse perspective algorithms showed its clear superiority in accuracy of road reconstructions.

3. Support of Martin Marietta

The University of Maryland played a critical support role in the development of Martin Marietta's visual navigation system. During the early months of the ALV project, Dr. Todd Kushner of our Laboratory (who had worked previously for the VICOM Corporation) provided valuable technical assistance to Martin Marietta in the use of the VICOM. Engineers from Martin Marietta spent several months at our Laboratory, studying computer vision and holding technical discussion with our staff on the design of visual navigation systems.

The first Martin Marietta demonstration in May 1985 used a version of a road inverse perspective algorithm designed by Dr. Allen Waxman. This algorithm, which is described in technical report [2], is capable of reconstructing road geometry including road banking.

Scientists from the University of Maryland also brought to Martin Marietta visual navigation software developed at Maryland for the VICOM on the ALV. In August 1986, Dr. Todd Kushner spent one month at Martin Marietta installing that software on the ALV, and used it to drive the ALV over portions of the test track. The details of this initial set of experiments are described in technical report [3]. In November of the following year, an expanded version of the software system was brought to Denver and installed on the ALV. It achieved higher operating speeds, and drove the ALV over longer segments of the test track. All of this software was made available to Martin Marietta.

The University of Maryland also organized a series of Vision Working Group meetings to discuss how the research community could best utilize the ALV and to design experiments that the community could run on the ALV. This group included representatives from SRI, Hughes, ADS, GE, Honeywell, Carnegie-Mellon and Martin Marietta. Based on the group's recommendations, several extensive data sets were collected from the ALV and distributed to interested members of the Strategic Computing Vision Technology Base.

4. Parallel Processing

Our Laboratory received three different parallel processing machines as part of its participation in the ALV project: a WARP systolic array processor, two Butterfly shared memory systems (one containing 128 MC68000 processors with one MB of memory per processor, and the second containing 16 MC68020 processors with 4 MB of memory per processor), and a 16K processor Connection Machine II. It has used all of these machines extensively to perform research in parallel vision for navigation. We discuss the use of each of these machines in the following subsections.

4.1. WARP

The WARP machine is ideally suited for low level and intermediate level vision algorithms. The first algorithm that we implemented on the WARP was the symmetric nearest neighbor image enhancement algorithm. This is an iterative image enhancement algorithm that replaces the grey level (or color) at each pixel in the image by the average of a subset of the pixel's neighbors, chosen in such a way as to ensure that those neighbors are most likely in the same image region as the pixel.

We also began work on a more significant research project using the WARP. The project involves range data processing, specifically for object (landmark) recognition. Several years ago one of our Ph.D. students (Dr. Teresa Silberberg) completed a thesis on three dimensional object recognition from TV images. She considered the special case in which the objects to be recognized were resting on a a plane whose orientation was known in the camera coordinate system. In this case it can be shown that the location of the object can be recovered just by matching two points in the image (say the images of object corners) to two points on the model surface. Since, a priori, it is difficult to decide which image features correspond to which model features, this basic location estimation procedure is embedded in a clustering algorithm. In 1987 we completed a study of how to modify this object recognition algorithm so that it can be applied to range data. A project to implement the algorithm was initiated during the contract period, but was not finished during the contract period.

4.2. Butterfly

The Butterfly was the first parallel processor installed in our Laboratory to support our research in visual navigation. The first project that we undertook on the Butterfly was the design and implementation of parallel algorithms for computing the Hough transform. The Hough transform was the algorithm employed by our visual navigation system to identify the locations of road boundaries in prediction windows; the computations associated with the Hough transform accounted for a large percentage of the time that the system spent performing image processing; parallel algorithms for this step could, thus, provide a speedup of the vision control loop. Technical report [5] contains a description of our Butterfly Hough transform algorithm. The algorithm was designed to provide good performance for a wide range of ratios of window size (i.e., number of pixels) to machine size (number of processors). Our algorithm achieved almost linear speedup over a wide range of problem sizes.

Next, we embedded this Hough transform algorithm into a complete road navigation system implemented on the Butterfly. This work was reported in Sunil Puri's Master's thesis, although never issued as a technical report. The Butterfly road navigation system introduced parallelism at a number of places, including window placement, road reconstruction and trajectory determination. Hardware communications problems prevented us from successfully using the Butterfly road navigation system to navigate our robot arm over the terrain board.

Finally, we used the Butterfly to study algorithms for parallel heuristic search. The results of this research were described in technical report [12] and the research continues under the current contract. We showed that the naive model for parallel search based on a shared OPEN list would quickly lead to saturation of the parallel processor, as processors bottlenecked on the critical section of removing work from or adding work to the OPEN list. The analysis was based on an adaptation of classical queuing theory models to shared memory computing.

4.3. Connection Machine

Finally, we have conducted a series of vision projects on the Connection Machine focused on its effective use for multiresolution vision and focus of attention vision. Here, our concern is with the efficient processing of images having far fewer pixels than there are processors in the Connection Machine. During the last year of the contract we developed two paradigms for processing small images efficiently, called fat images and replicated images.

In a fat image we utilize many CM processors to represent a single pixel in the image. So, for example, if we are processing a 32×32 image on a 16K Connection Machine, then we would allocate 16 processors per pixel. These processors are simultaneously utilized by distributing the bits representing a pixel's grey level across the processors. In technical report [9] we describe how to implement the basic image processing operations of histogramming, table lookup, arithmetic and convolution using the fat image.

Unfortunately, our Connection Machine implementations of the fat image processing algorithms revealed that the Connection Machine was not well suited for this type of processing. Therefore, we began a study of an alternative processing strategy, which we call replicated image processing. Here, if we have k times as many processors as pixels we store k complete copies of the image in the Connection Machine. Towards the end of the contract period we had completed the design of all the basic image processing algorithms for replicated images, and are now implementing these algorithms under the current contract.

5. Conclusions

Before the introduction of the ALV program there had been little research, especially experimental research, conducted on the problem of visual navigation. One of the important scientific goals of the ALV project was to develop within the research community a conceptual framework for the study of visual navigation systems. While this goal was not completely fulfilled within the first three years of the program, significant progress was made in identifying the key theoretical and experimental problems that should be addressed by the community over the next several years.

Perhaps the most important is the integration of vision and planning into a unified framework. Early in the ALV program there were a series of informal meetings held between representatives of the planning community and the vision community with the (retrospectively) naive goal of identifying the "interface" between planning and vision for the specific problem of road navigation. It eventually became clear that the planning models available at that time were inadequate for two principle reasons:

1) They all depended on the availability of a level of representation of the world that was both more accurate and more abstract than one could reasonably hope to obtain with current perceptual systems.
2) There was no framework for planning that effectively combined reactive planning (i.e., the ability to respond to temporally unpredictable external events) and classical static plan generation systems.

While progress has been made in the planning community during the past three years on the second problem, it has mostly proceeded independently of the considerations of (1).

Intimately related to the problem of integrating vision and planning is the identification of appropriate control level architectures for a visual navigation system. The road navigation system developed at Maryland had a very classical control architecture, with modules for image processing, image prediction, sensor control, modest geometric reasoning, path planning and path execution. These were all controlled by a so-called vision executive that routed relevant information between the modules. With the exception of the "evolutionary" architecture proposed by Brooks [21] at MIT, it seems that most visual navigation systems have been designed along lines similar to the Maryland system. It is not clear that they are adequate for developing systems pursuing many navigation goals simultaneously (e.g., maintaining visual stabilization while moving towards some target). These problems should receive some attention over the next several years.

A third important issue that arose during the course of our research was the extent to which autonomous visual navigation systems have to be based on a reconstructive approach to vision as opposed to an associative approach. This dichotomy was discussed at some length in Randall Nelson's Ph.D. thesis (supported by our DARPA Image Understanding Project) [22]. The reconstructive approach involves using vision to construct a three dimensional representation of those parts of the visual environment needed to perform some navigation task. So, for example, our road navigation system operated by reconstructing the three dimensional geometry of the road boundaries. At the other extreme, an associative approach would specify a direct relationship between uninterpreted image properties (such as statistics of edge direction distributions) and navigation actions. The visual homing system developed by Randal Nelson as part of his Ph.D. thesis operated by using reduced resolution edge maps as an index into a large associative table of motor control commands. An interesting intermediate approach would involve the identification of image structures that have three dimensional significance, without necessarily completely determining their three dimensional structure. So, for example, one can construct a road following system that identifies the road boundaries in an image, and then issues motor control commands based on their image locations and orientations. While there would be some implicit three dimensional model underlying the determination of the motor controls, the system itself would not operate by explicitly reconstructing the three dimensional geometry of the road. It might be the case that visual navigation systems have to operate using all three of these models at different times based on the current task set. Further research over the next several years will hopefully shed some light on these issues also.

6. Relevant Technical Literature

1. Jacqueline Le Moigne, Allen M. Waxman, Babu Srinivasan and Matti Pietikainen, "Image Processing for Visual Navigation of Roadways." CAR-TR-138, CS-TR-1536, DACA76-84-C-0004, July 1985.

ABSTRACT: A system which supports the visual navigation of roadways by an autonomous land vehicle has been developed at the Computer Vision Laboratory. One of the modules of this system is an Image Processing Module which extracts 2-D symbolics from the imagery to be analyzed in the world domain by Reasoning and Geometry Modules. In this report, we describe the Image Processing Module. Different representations can be used in the image domain: boundary-based and region-based are two examples of such representations. We present two kinds of algorithms for extracting roads from imagery, corresponding to these two different representations: linear feature extraction and gray-level or color segmentation. For each kind of processing, two different modes, called "bootstrap" and "feed-forward", may be utilized. The bootstrap mode processes the entire image and assumes no prior information about the location of a road, while the feed-forward mode utilizes a prediction derived from the processing of a prior road segment and the distance traversed in order to focus visual attention. These algorithms are described in detail, and example results are shown. The examples include real road images and "simulator images" obtained in our laboratory with a scale model system comprised of a road network and a robot arm carrying a camera.

2. Allen M. Waxman, Jacqueline Le Moigne, Larry S. Davis, Eli Liang and Tharakesh Siddalingaiah, "A Visual Navigation System for Autonomous Land Vehicles." CAR-TR-139, CS-TR-1537, DACA76-84-C-0004, July 1985.

ABSTRACT: A modular system architecture has been developed to support visual navigation by an autonomous land vehicle. The system consists of vision modules performing image processing, 3-D shape recovery, and rule-based reasoning, as well as modules for planning, navigating and piloting. The system runs in two distinct modes, *bootstrap* and *feed-forward*. The bootstrap mode requires analysis of entire images in order to find and model the objects of interest in the scene (e.g., roads). In the feed-forward mode (while the vehicle is moving), attention is focused on small parts of the visual field as determined by prior views of the scene, in order to continue to track and model the objects of interest. We have decomposed general navigational tasks into three categories, all of which contribute to planning a vehicle path. They are called *long, intermediate, and short range navigation*, reflecting the scale to which they apply. We have implemented the system as a set of concurrent, communicating modules and use it to drive a camera (carried by a robot arm) over a scale model road network on a terrain board.

3. Larry S. Davis and Todd Kushner, "Road Boundary Detection for Autonomous Vehicle Navigation." CAR-TR-140, CS-TR-1538, DACA76-84-C-0004, July 1985.

ABSTRACT: The Computer Vision Laboratory at the University of Maryland has been participating in DARPA's Strategic Computing Program for the past year. Specifically, we have been developing a computer vision system for autonomous ground navigation of roads and road networks. The complete system runs on a VAX 11/785, but certain parts of the system have been reimplemented on a VICOM image processing system for experimentation on an autonomous vehicle built for the Martin Marietta Corp., Aerospace Division in Denver, Colorado. We give a brief overview here of the principal software components of the system, and then describe the VICOM implementation in detail.

4. Daniel DeMenthon, "Inverse Perspective of a Road from a Single Image." CAR-TR-210, CS-TR-1685, DACA76-84-C-0004, July 1986.

ABSTRACT: A method is presented for reconstructing the geometry of a road from a single image of the road. This problem has an infinity of solutions unless restrictive hypotheses about geometric characteristics of the road are assumed. The road is modeled as a space ribbon defined by a spine (centerline) and generators which are horizontal line segments cutting the spine at their midpoint at a right angle. Properties of two neighboring generators of such a ribbon are examined, and it is found that if a generator is known, a neighbor is completely defined if one of its ends is known. The proposed method uses this property to reconstruct the visible part of the world road, by iteratively finding a series of generators. For validation of this method, a road making an S on a hill or in a valley is defined analytically and graphically, and its perspective image is obtained; from this image, algorithms reconstruct a ribbon which is compared to the original world model by superposition, in top view and side view. The proposed method is tested against a simple method which assumes that the ground is flat ("flat earth assumption"), and against another method which uses vanishing points.

5. Sharat Chandran and Larry S. Davis, "The Hough Transform on the Butterfly and the NCUBE." CAR-TR-226, CS-TR-1713, DACA76-84-C-0004, September 1986.

ABSTRACT: This report describes the parallel implementation of the Hough Transform, a technique to detect collinear edge points. Specifically, we consider two contrasting architectures, the Butterfly Parallel Processor[1], essentially a shared memory machine, and the NCUBE[2], a direct connection machine in which processors are interconnected in the form of a hypercube.

[1]Butterfly is a trademark of Bolt Beranek and Newman, Inc.

[2]NCUBE is a trademark of NCUBE Corporation.

Developing parallel Hough transform algorithms involves addressing questions of optimal processor allocation and parallel "peak" selection in image neighborhoods. We present fast practical algorithms (subject to inherent lower bounds) and discuss the relevant complexity issues.

6. Jacqueline Le Moigne, "Domain-Dependent Reasoning for Visual Navigation of Roadways." CAR-TR-230, CS-TR-1721, DACA76-84-C-0004, October 1986.

ABSTRACT: A Visual Navigation System for Autonomous Land Vehicles has been designed at the Computer Vision Laboratory of the University of Maryland. This system includes several modules, among them a "Knowledge-based Reasoning Module" that is described in this report. This module utilizes domain-dependent knowledge (in this case, "road knowledge") in order to analyze and label the visual features extracted from the imagery by the Image Processing Module. Knowledge and general hypotheses are given in Section 2. The Reasoning Module itself is described in Section 3 and results are presented in Section 4. Finally, some conclusions and future extensions are proposed in Section 5.

7. Sunil Puri and Larry S. Davis, "Two Dimensional Path Planning with Obstacles and Shadows." CAR-TR-255, CS-TR-1760, DACA76-84-C-0004, January 1987.

ABSTRACT: A mobile robot navigates with a limited knowledge of its environment because of the restricted field of view and range of its sensors, and the occlusion of parts of the environment in any single image. Most path planning algorithms consider only free regions and obstacles in the robot's world for path planning. The objective of this report is to extend the classical path planning paradigm to include occluded regions. This introduces the new problem of deciding when (or whether) to employ the sensor system during the execution of the path to, potentially, reveal the occluded regions as obstacles or free space for the purpose of replanning.

8. Minoru Asada and Saburo Tsuji, "Shape from Projecting a Stripe Pattern." CAR-TR-263, CS-TR-1773, DACA76-84-C-0004, January 1987.

ABSTRACT: This paper presents a simple method which determines the shape of an object by projecting a stripe pattern on to it. Assuming orthographical projection as a camera model and parallel light projection of the stripe pattern, the method obtains a 2(1/2)D representation of objects by estimating surface normals from the slopes and intervals of the stripes in the image. The 2(1/2)D image is further divided into planar or singly curved surfaces by examining the distribution of the surface normals in gradient space. Some applications and evaluation of the error in surface orientation are described.

9. Thor Bestul and Larry S. Davis, "On Computing Histograms of Images in $\log n$ Time Using Fat Pyramids." CAR-TR-271, CS-TR-1791, DACA76-84-C-0004, February 1987.

ABSTRACT: This paper presents an algorithm for the $\log n$ computation of the complete histogram of an $n \times n$ gray-level image. It uses a "fat" pyramid implemented on an SIMD hypercube multiprocessor with very high processor utilization. A "fat" pyramid is a pyramid in which the size of a processor associated with a node in the pyramid depends on the level of the pyramid in which the node appears. We describe how to embed fat pyramids in hypercubes using Gray codes, and then describe the histogramming algorithm.

10. Sharat Chandran, Larry S. Davis, Daniel DeMenthon, Sven J. Dickinson, Suresh Gajulapalli, Shie-Rei Huang, Todd R. Kushner, Jacqueline Le Moigne, Sunil Puri, Tharakesh Siddalingaiah and Phillip Veatch, "An Overview of Vision-Based Navigation for Autonomous Land Vehicles 1986." CAR-TR-285, CS-TR-1831, DACA76-84-C-0004, April 1987.

ABSTRACT: This report describes research performed during the first two years on the project *Vision-Based Navigation for Autonomous Land Vehicles* being conducted under DARPA support. The report contains discussion of four main topics:

1) Development of a vision system for autonomous navigation of roads and road networks.
2) Support of Martin Marietta Aerospace, Denver, the integrating contractor on DARPA's ALV program.
3) Experimenting with the vision system developed at Maryland on the Martin Marietta ALV.
4) Development and implementation of parallel algorithms for visual navigation on the parallel computers developed under the DARPA Strategic Computing program—specifically, the WARP systolic array processor, the Butterfly and the Connection Machine.

11. Minoru Asada, "3-D Road Structure from Motion Stereo." CAR-TR-286, CS-TR-1839, DACA76-84-C-0004, April 1987.

ABSTRACT: This paper presents a new method for reconstructing the 3-D structure of road boundaries from consecutive images. First, we present a method for estimating depth information by applying a motion stereo method to consecutive images, given an estimate of the interframe motion. The relation between depth, motion and disparity is investigated, since the accuracy of the depth depends on the disparity range. Next, the error of the estimated road structure due to quantization errors and motion estimation errors is examined. Finally, a representation for road boundaries is proposed that makes explicit the error of the road edge location in 3-D space. Experimental results are shown for an input image sequence taken by the ALV simulator robot in the Center for

Automation Research at the University of Maryland.

12. Shie-rei Huang and Larry S. Davis, "A Tight Upper Bound for the Speedup of Parallel Best-First Branch-and-Bound Algorithms." CAR-TR-290, CS-TR-1852, DACA76-84-C-0004, May 1987.

ABSTRACT: Most previous studies of the speedup of parallel branch-and-bound algorithms are based on the amount of work done in the parallel case and in the sequential case[14,17,18,23]. Any evaluation of a parallel algorithm should include both the execution time and the synchronization delay[30]. In this paper, a finite population queueing model is used to capture the synchronization delay in parallel branch-and-bound algorithms and to quantitatively predict the behavior of their speedup. A program to solve the Traveling Salesman Problem was written on a BBN Butterfly[2] multiprocessor to empirically demonstrate the credibility of this theoretical analysis. Finally, we note that similar analyses can be applied to evaluate parallel AI systems in which processes communicate through a shared global database.

13. Phillip A. Veatch and Larry S. Davis, "Range Imagery Algorithms for the Detection of Obstacles by Autonomous Vehicles." CAR-TR-309, CS-TR-1888, DACA76-84-C-0004, July 1987.

ABSTRACT: Algorithms are presented which segment range images and classify regions as being navigable or unnavigable by a land vehicle. The algorithms are applied to data collected from an active laser range sensor mounted on an autonomous land vehicle and their comparative results are analyzed.

The sensitivity of various algorithms to uncertainty in the orientation of the range sensor is studied. Experiments on sensor calibration and image enhancement are also presented.

A computer model of an autonomous land vehicle and its environment is described which provides a valuable tool for investigating many issues of navigation with range sensors. Obstacle detection algorithms are used in conjunction with the model to demonstrate a vehicle navigating itself through an obstacle strewn world to a goal location.

14. Phillip A. Veatch and Larry S. Davis, "IRS: A Simulator for Autonomous Land Vehicle Navigation." CAR-TR-310, CS-TR-1889, DACA76-84-C-0004, July 1987.

ABSTRACT: IRS is a computer simulation program that provides a software testbed for autonomous navigation algorithms. The program allows the user to describe a complex world built from spheres, parallelepipeds, planar surfaces, cones, and cylinders. The program simulates the movement of an Autonomous Land Vehicle and constructs video and range images based on the ALV's field of view as the vehicle moves through the world. Ground maps of the world, as perceived by the ALV, are also created.

15. Sven J. Dickinson and Larry S. Davis, "An Expert Vision System for Autonomous Land Vehicle Road Following." CAR-TR-33, CS-TR-1932, DACA76-84-C-0004, October 1987.

ABSTRACT: A production system model of problem solving is applied to the design of a vision system by which an autonomous land vehicle (ALV) navigates roads. The ALV vision task consists of hypothesizing objects in a scene model and verifying these hypotheses using the vehicle's sensors. Object hypothesis generation is based on the local navigation task, an a priori road map, and the contents of the scene model. Verification of an object hypothesis involves directing the sensors toward the expected location of the object, collecting evidence in support of the object, and reasoning about the evidence. Constructing the scene model consists of building a semantic network of object frames exhibiting component, spatial, and inheritance relationships. The control structure is provided by a set of communicating production systems implementing a structured blackboard; each production system contains rules for defining the attributes of a particular class of object frame. The combination of production system and object oriented programming techniques results in a flexible control structure able to accommodate new object classes, reasoning strategies, vehicle sensors, and image analysis techniques.

16. Minoru Asada, "Building a 3-D World Model for a Mobile Robot from Sensory Data." CAR-TR-332, CS-TR-1936, DACA76-84-C-0004, October 1987.

ABSTRACT: This paper presents a method for building a 3-D world model for a mobile robot from sensory data. The 3-D world model consists of three kinds of maps: a sensor map, a local map and a global map. A range image (sensor map) is transformed to a height map (local map) with respect to a mobile robot. First, the height map is segmented into four categories (unexplored, occluded, traversable, and obstacle regions) for obstacle detection and path planning. Next, obstacle regions are classified into artificial objects (buildings, cars, road signs, etc.) or natural objects (trees, bushes, etc.) using both the height image and video image. One drawback of the height map—the recovery of vertical planes—is overcome by the utilization of multiple height maps which include the maximum and minimum height of each point, and the number of points in the range image mapped into one point in the height map. The multiple height map is useful not only for finding vertical planes in the height map but also for segmentation of the video image. Finally, the height maps are integrated into a global map by matching geometrical properties and updating region labels.

The method is tested on a model including many objects such as trees, buildings, cars, and so on.

17. Daniel DeMenthon, Tharakesh Siddalingaiah and Larry S. Davis, "Production of Dense Range Images with the CVL Light-Stripe Range Scanner." CAR-TR-337, CS-TR-1962, DACA76-84-C-0004, December 1987.

ABSTRACT: This report describes a system able to produce 512×512 range images of model scenes in the laboratory. This ranging instrument, which comprises a light-emitting slit, a cylindrical lens, a step-motor controlled mirror and a CCD camera, is compact enough to be mounted on the tool plate of a robot arm. The light source itself is mounted away from this structure, and the light is brought to the slit by a flexible fiberoptic light guide. The robot arm's motion can be controlled by inputs from the range scanner, for simulation of autonomous vehicles equipped with rangers. This system is programmed to produce range images which are comparable in many respects to range images produced by laser range scanners. With this similitude of formats, software for edge detection, object recognition, dynamic path planning or data fusion with video images can be developed on range images produced by this laboratory equipment and can be easily ported to laser ranging systems.

18. Larry S. Davis, "Vision-Based Navigation for Autonomous Ground Vehicles—First Annual Report." AD-A203 712, DACA76-84-C-0004, July 1988.

19. Larry S. Davis, "Vision-Based Navigation for Autonomous Ground Vehicles—1986 Annual Report." AD-A207 596, DACA76-84-C-0004, August 1986.

20. Larry S. Davis, "Vision-Based Navigation for Autonomous Ground Vehicles—Third Annual Report." AD-A171618, DACA76-84-C-0004, November 1988.

ABSTRACT: This is the third annual report for DARPA Contract DACA76-84-C-0004 (DARPA Order 5096), covering the period July 1986 through July 1987. The report describes both new equipment added to our laboratory and the research performed on autonomous vehicle navigation. We describe the design of a structured light range scanner that has been built and mounted on our robot arm. This scanner provides us with the capability of generating range data similar to that obtainable on the ALV using the ERIM scanner. The report also describes the following research projects conducted during the past year:

1) The design and implementation of a rule-based road following system. This system has provided us with a flexible environment in which to experiment with different visual control strategies for road extraction.

2) Road obstacle detection in range data. We have developed computationally simple algorithms for road obstacle detection and applied them to a variety of synthetic and real range imagery. Simple geometric arguments show why these algorithms should be more robust than those used currently on the

ALV to detect obstacles.

3) Theoretical analysis of the accuracy of road recovery using motion stereo. Here, our research shows that it is unlikely that the three-dimensional structure of the road can be recovered with sufficient accuracy from motion stereo, given the expected errors in the estimate of vehicle motion.

4) Parallel vision on the Connection Machine. Here, we introduce a computational structure called a Fat Pyramid, and show how the common operation of histogramming can be implemented within the fat pyramid structure. Fat Pyramids provide a possible means for effectively utilizing the Connection Machine hardware for either multiresolution or focus of attention vision algorithms.

Finally, the report ends with a discussion of our plans for research during the next three years of the ALV program.

21. R. Brooks, "A robust control system for a mobile robot," *IEEET-Robotics and Automation* **2**, 14–23.

22. Randal C. Nelson, "Visual Navigation." CAR-TR-380, CS-TR-2087, DAAB07-86-K-F073, August 1988.

ABSTRACT: Visual navigation is a major goal in machine vision research, and one of both practical and basic scientific significance. The practical interest reflects a desire to produce systems which move about the world with some degree of autonomy. The scientific interest arises from the fact that navigation seems to be one of the primary functions of vision in biological systems. Navigation has typically been approached through reconstructive techniques since a quantitative description of the environment allows well understood geometric principles to be used to determine a course. However, reconstructive vision has had limited success in extracting accurate information from real-world images. This report argues that a number of basic navigational operations can be realized using qualitative methods based on inexact measurement and pattern recognition techniques.

Navigational capabilities form a natural hierarchy beginning with simple abilities such as orientation and obstacle avoidance, and extending to more complex ones such as target pursuit and homing. Within a system, the levels can operate more or less independently, with only occasional interaction necessary. This report considers three basic navigational abilities: ***passive navigation***, ***obstacle avoidance***, and ***visual homing***, which together represent a solid set of elementary, navigational tools for practical applications. It is demonstrated that all three can be approached by qualitative, pattern-recognition techniques. For passive navigation, global patterns in the spherical motion field are used to robustly determine the motion parameters. For obstacle avoidance, divergence-like measurements on the motion field are used to warn of potential collisions. For visual homing an associative memory is used to construct a system which can be trained to home visually in a wide variety of natural environments. Theoretical analyses of the techniques are presented, and implementation and testing of working systems described.